Available online at www.sciencedirect.com
ScienceDirect
Procedia Economics and Finance 12 (2014) 771 – 779
Procedia Economics and Finance
www.elsevier.com/locate/procedia

Enterprise and the Competitive Environment 2014 conference, ECE 2014, 6–7 March 2014, Brno, Czech Republic

Quality of Experience of Voice Services in Corporate Network

Petr Zach[a,]*, Martin Pokorny[a], Jiri Balej[a]

[a]*Department of Informatics FBE MENDELU in Brno, Zemedelska 1, 61300 Brno, Czech Republic*

Abstract

This paper aims on a quality estimation of voice services in converged corporate networks. Besides the techno-centric Quality of Service (QoS) metrics a research area called Quality of Experience (QoE) provides metrics and methods for quality evaluation from the end-user's perspective. This contribution focuses on a QoE estimation of Voice over IP (VoIP) calls. Existing methods of voice quality estimation are compared on different voice codecs tested on a network topology suffering from distortions of real network. Finally, a regression analysis is employed to provide better understanding of impact of network conditions (delay, jitter, packet loss) on the VoIP service quality. The results can improve the voice quality monitoring systems in corporate networks.

Selection and/or peer-review under responsibility of the Organizing Committee of ECE 2014

Keywords: Voice; VoIP; quality of service; quality of experience; MOS; computer network

1. Introduction

The demand for the voice services delivered by computer network has been increased recently. With the rapid development of mobile devices and almost ubiquitous wireless connection, the spectrum of services in computer networks changes. Besides the traditional data, more voice and video services traverse the converged networks. As user requirements rise, the conditions of the voice and video services are stricter. Consequently, the amount of carried data is often larger than the physical resources along the connection can accommodate. This can cause several data impairments. Moreover, the multimedia services are significantly more sensitive to network distortions in comparison with traditional data (e.g. web browsing). Once a temporal network distortion appears, the toolset

* Corresponding author. Tel.: +420-545-132-237.
E-mail address: petr.zach@mendelu.cz

Selection and/or peer-review under responsibility of the Organizing Committee of ECE 2014
doi:10.1016/S2212-5671(14)00404-3

called Quality of Service (QoS) should help multimedia services to overcome this impairment but it may not always be sufficient.

Since the quality of service is not guaranteed in the current Internet, it is important for the QoS to be constantly monitored so that the applications can take proper actions needed to ensure appropriate service quality level to meet the end-user requirements. The low quality perception has i.a. direct impact on the profit of these services as proofed by Klein and Jakopin (2014). There are three basic quantitative QoS metrics (delay, jitter, packet loss) used for the end-to-end quality measurement. Besides to the QoS, a Quality of Experience (QoE) research area provides qualitative metrics and methods that better reflect the end-user perception of the delivered service. The main goal of QoE research area is to find relations and their weights between QoS and QoE metrics to accurately estimate the final service quality. This article focuses on VoIP (Voice over IP) service that is commonly used as a voice service in the corporate environments. The results of this contribution can improve network monitoring systems by real-time estimation of the final service quality degree.

1.1. Objectives

The aim of this article is to observe an impact of different computer network distortions on the voice service quality, and propose a new statistical model useful for the real-time voice service quality evaluation. First, the quality estimation problem is formulated and related work reviewed (in sections 2 and 3). After that, the experiment methodology is introduced in section 4. The results (in section 5) describe the experiment and the proposed statistical model. Conclusion and future work follows in section 6.

2. Problem formulation

The term QoS is usually understood very differently. Its meaning ranges from the intrinsic network parameters to the user perceivable features of the service and user experience. In fact, the user experience is covered by QoE. Stankiewicz and Jajszczyk (2011) describe the intrinsic QoS as a network performance or as a set of parameters containing delay, jitter, packet loss, error probability, available bandwidth, etc. Not only these parameters influence the resulting QoE, moreover their mutual relationships differ with the type of service and the end-user expectations. Kuipers et al. (2010) describe the QoE as overall acceptability of an application or service, as perceived subjectively by the end-user. This definition refers to QoE as an area of subjective (qualitative) measurements. QoE introduces a Mean Opinion Score (MOS) to express the final service quality level. ITU-T (1996) defines MOS as a five point scale where 5 means excellent, 4 good, 3 fair, 2 poor and 1 is bad. Lozano-Garzon (2012), Stankiewicz and Jajszczyk (2011) describe two approaches for the QoE measurement: subjective (qualitative) and objective (quantitative).

2.1. Subjective measurement

The subjective methods are built with the participation of people, a representative sample of the population that uses the particular service. In these methods the service is evaluated in a controlled environment and respondents fill out a survey with numerical values. Defined by ITU-T (1998) in recommendation P.800, the Absolute Category Rating (ACR) test is one of the widely accepted norms for subjective speech quality rating.

2.2. Objective measurement

The results of the subjective QoE assessment are usually excellent but it is technically and economically impossible to hire a group of people to examine service quality for the each particular measurement. Under these conditions, the researchers tend to propose the objective technics. The objective methods provide the QoE assessment based on the measurement of several technical (QoS) parameters less or more related to the service quality. Ding (2007) and Silva (2008) classify the objective methods into two categories: intrusive (reference) and non-intrusive (no-reference).

In the intrusive methods, the MOS is measured by comparing the reference signal with the degraded one, which is the output of the system under test. Intrusive methods can achieve relatively accurate MOS estimates. However, they are not suitable for the real-time, live call quality monitoring purposes, as in this case the reference speech is unavailable. Currently, the PESQ, defined by ITU-T (2001), and the POLQA, defined by ITU-T (2011), standards are widely used in the industry. Expect of them, the AQuA (Sevana Oy, 2014) method was recently developed to extend the range of available methods. According to Silva (2008), these metrics are geared toward listening quality, not on conversational quality. This is an important issue, since listening quality only concerns the actual sound quality of the speech stream, whereas the conversational quality refers to the overall quality of the conversation, including interactivity aspects.

The non-intrusive methods utilize the degraded speech signal only or estimate the final QoE degree on some statistics collected from the network. The well-known example of the no-reference model is the E-model defined by ITU-T (2005). The E-model estimates the end-users experience during a voice conversation based on the end-device characteristics and the transport parameters. Moreover, another approaches based on random neural network (RNN) known as PSQA (Pseudo Subjective Quality Assessment) methods are a part of the non-intrusive QoE estimation methods. The estimation accuracy of the non-intrusive methods is lower but these tools are very usable for real-time quality estimation.

3. Related work

The area of QoE including factors influencing the service perception in converged networks is thoroughly described by Stankiewicz and Jajszczyk (2011). Fixed networks and wireless networks based on different technologies are considered. A variety of technologies and concepts for future converged networks are discussed.

Ding et al. (2007) propose a parametric, non-intrusive speech quality assessment algorithm suitable for VoIP environments. The voice payload analysis is performed by the PESQ algorithm and the noise perception model is incorporated in the ITU-T E-model. The quality of experience (QoE) of Skype calls over UMTS was measured and analyzed by Hoßfeld and Binzenhöfer (2008) in terms of the MOS value. Additionally, the classic QoS parameters like throughput or jitter are measured to derive a traffic profile for the Skype application. Requera, et al. (2008) apply non-intrusive E-model quality estimation to assess the impact of active queue management (AQM) on the quality of service of voice over IP (VoIP). Some of the most representative AQM schemes are analyzed through extensive simulations. Besides the statistical non-intrusive methods, there are plenty of articles focusing on the application of neural networks usable for the QoE quality estimation, e. g. Škorpil and Šťastný (2006). Authors Lozano-Garzo et al. (2012) propose a RNN as a part of PSQA method used for the web service quality estimation.

The field of QoE covers not only the area of voice services but investigates the perceived quality of video and in last years the web quality as well. Alvares et al. (2013) and Santos et al. (2014) deal with the QoE of video services where voice and video must be evaluated simultaneously and where the subjective evaluation is mostly necessary. Inacio and Cruz (2013) observe the impact of several attributes of streamed video (resolution, video theme) on the perceived video quality, finally a statistical model is proposed. On the QoE of web traffic focus Papageorgiou et al. (2013).

The understanding of QoE helps to monitor and actively influence the data flows in the network to preserve the required quality during the network impartments or the network congestion. Bayer (2010), Taboada et al. (2013) or Seppänen et al. (2014) propose QoS/QoE aware frameworks for computer network flows management.

As evident from the previous sections, the QoE is an interesting and still developing area of various approaches focusing on quality measurement of all types of services in converged networks. The vast majority of subjective and objective measurements follow ITU-T recommendations and methods (ACR, E-model, PESQ, etc.). No publication found dealing with the VoIP service quality estimation using algorithm AQuA.

4. Experiment methodology

The goal of this experiment is to examine the impact of delay, jitter and packet loss on perceived quality of selected voice codecs (G.711 and G.729). Besides the well-known approaches for the objective service quality evaluation, the AQuA software was used. Using MOS values calculated by this tool a new statistical model for the voice codecs quality evaluation will be proposed.

4.1. Test topology

The network topology shown on the topology diagram in Fig. 1 consists of two 1 Gbps local networks separated by a router. Each of these networks contains one computer (2 GB RAM, 1 CPU) where Windows 7 64 bit is running. The computer on the left side acts as a VoIP call sender (talking side), the computer on the right side operates as a VoIP call receiver (listening side). The router is running Linux CentOS 6.3 64 bit (GB RAM, 1 CPU). As shown in the Fig. 1, the router passes packets between its eth0 and eth1 interfaces. While the direction from the right to the left has no limits, some network distortions are intentionally involving packets leaving eth1 (direction from left to right) to simulate network impairments common in WAN networks. These all devices were virtualized by VirtualBox 4.3.6 in the Windows 8 64 bit host system (CPU Intel Core i5-3437U, 8 GB RAM).

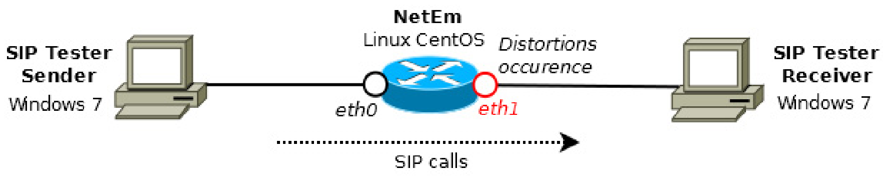

Fig. 1. The experiment topology diagram.

4.2. Software

The NetEm utility, a part of Linux kernel, was used as a network emulator. Several network distortions, e.g. delay, jitter, packet loss, can be placed in the network using this tool. The tested VoIP calls were generated and recorded by Trinity SIP Tester (StarTrinity, 2014). The sender initiates a VoIP call and plays a predefined sound. The receiver records and stores the receiving voice. Finally, the voice quality evaluation is performed by the AQuA.

4.3. VoIP calls

Two commonly used voice codecs were examined – G.711 (a-law) and G.729. The details of these codecs are described in Tab. 1, including required bandwidth, the MOS during the optimal conditions and the compression type. The SIP protocol was used as a signaling protocol. The voice data were encapsulated into the RTP (Real-time Transport Protocol), to the UDP respectively. The transmitted voice was represented by an instrumental music record in WAV (16 bit PCM) format. Compared with the speech record this one has no blank spaces in its waveform to better recognize each quality decrease across the whole record.

Table 1. The voice codecs attributes.

Description	Bandwidth	MOS	Compression
G.711	64 kbps	4.1	Lossless
G.729	8 kbps	3.92	Lossy

4.4. Test scenario

The test scenario consists of two phases. Voice quality coded by G.711 was examined during the first phase followed by G.729 in the second phase. Each phase consists of a set of particular voice quality measurements where the different network distortions were applied during the VoIP call. Afterwards, the AQuA software was used to gain measured MOS values (MOS_M) of the particular VoIP call. Values of delay, jitter and packet loss were systematically changed to test the listening quality of the voice record (VoIP call). In this scenario, delay was between 0 a 500 ms (increased by 10 ms in each step), jitter was between 0 to 100 ms (increased by 5 ms in each step), packet loss was between 0 and 10 % and was increased by 0.5 percentage point in every step.

5. Results

The results obtained from our experiments are described in this section. We do so by showing how different parameters affect the perceived listening quality, according to the AQuA quality assessments. After the simulation process, obtained MOS_M values were collected and the Pearson correlation coefficient (PC) calculated. As seen in Tab. 2., there is a strong negative correlation between the jitter and the MOS_M values in case of both of the voice codecs. On the other hand, the delay has negligible impact on the MOS_M, thus the delay will not be considered in further work. Packet loss has considerable impact on the MOS_M only in case of the G.711 and will be part of further analysis, even though its impact on the MOS_M is minimal in case of the G.729.

Table 2. The Pearson Correlation Coefficient of the MOS_M and particular QoS metrics.

Description	Delay	Jitter	Packet Loss
G.711	−0.0345	−0.8172	−0.2294
G.729	−0.0189	−0.8322	−0.0237

5.1. MOS as a function of one QoS metric

Assuming that only one QoS parameter (jitter or packet loss) will be taken into account, the MOS of the G.711 voice quality can be according our results expressed as a function of jitter (J) by linear regression (Least Squares Method – LSM) described in equation 1 with R^2=0.7218. Alternatively, the MOS of the G.711 voice quality can be expressed as a function of packet loss (L) by linear regression described in equation 2, R^2=0.526. Plot in Fig. 2a, resp. in Fig. 2b, shows how jitter impacts the measured quality (MOS_J), resp. how packet loss impacts the measured quality (MOS_L), of the G.711 VoIP call.

$$MOS_J^{G.711} = 3.5086 \times e^{-0.013 \times J} \tag{1}$$

$$MOS_L^{G.711} = -0.0785 \times L + 2.4192 \tag{2}$$

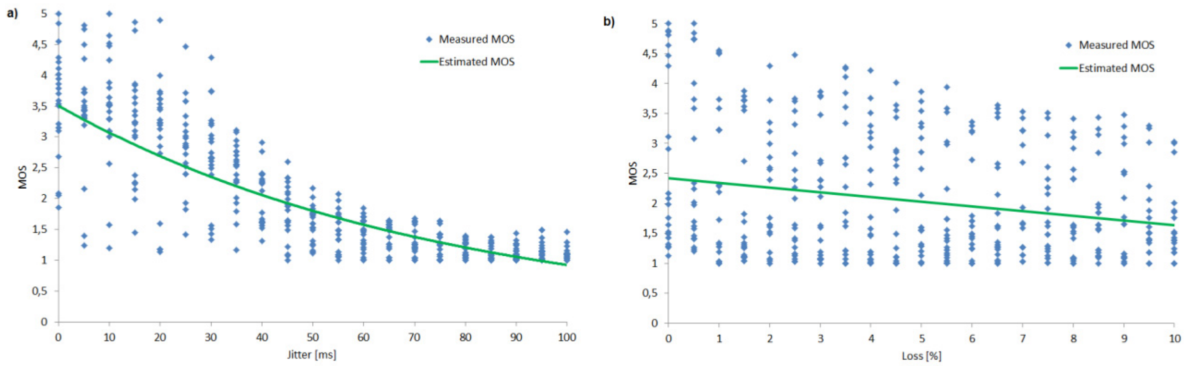

Fig. 2. (a) The impact of jitter on the MOS of the G.711; (b) the impact of packet loss on the MOS of the G.711.

Because of the lossy compression of the G.729 codec, the absolute MOS_M values were noticeably lower in comparison to the G.711 VoIP calls. Despite this fact, the MOS_M values from interval between 0 and 20 ms of jitter are slightly lower than expected. The reason remains an open question for us. Anyway, the listening voice quality can be expressed as a function of jitter (J) by linear regression described in equation 3 with R^2=0.7251. Alternatively, the MOS of the G.729 voice quality can be expressed as a function of packet loss (L) by linear regression described in equation 4, R^2=0.058. Plot in Fig. 3a, resp. in Fig. 3b, shows how jitter impacts the measured quality (MOS_J), resp. how packet loss impacts the measured quality (MOS_L), of the G.729 VoIP call.

$$MOS_J^{G.729} = 2.7153 \times e^{-0.01 \times J} \tag{3}$$

$$MOS_L^{G.729} = -0.0154 \times L + 1.857 \tag{4}$$

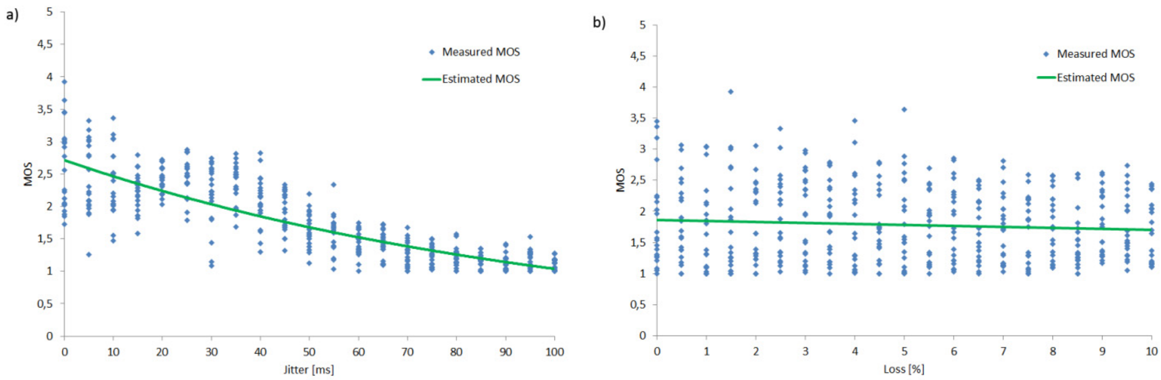

Fig. 3. (a) The impact of jitter on the MOS of the G.729; (b) the impact of packet loss on the MOS of the G.729.

5.2. MOS as a function of jitter and packet loss

In the previous section, different formulas were presented to characterize the behavior of MOS results where different particular QoS parameters were taken into account separately. This section describes the mathematical procedure how to encapsulate the different formulas from the text above into a single and global function. It is necessary to combine previous functions to accurately predict (estimate) the MOS values (MOS_E). This combination will follow multiple linear regression model represented in equation 5. The estimated MOS_E value can be seen as the dependent variable, modeled as a function of the MOS_J and MOS_L values and their corresponding linear weights where β_0 is the offset value, α_i is the linear weight associated to each QoS parameter, x_i represents MOS_J, resp. MOS_L.

$$MOS_E = \beta_0 + \sum_{i=1}^{n} (\alpha_i \times x_i) \tag{5}$$

Regarding the experiment results, the global prediction function is achieved and is represented by equation 6 in case of the G.711 VoIP calls and by equation 7 in case of the G.729 VoIP calls. The graphical representations of proposed models are shown in Fig. 4b and in Fig. 5b. The original MOS values (MOS_M) are depicted in Fig. 4a and in Fig. 5a.

$$MOS_E^{G.711} = -2.24289 + 1.13468 \times MOS_J^{G.711} + 0.999917 \times MOS_L^{G.711} \tag{6}$$

$$MOS_E^{G.729} = 5.16043 + 0.838930 \times MOS_J^{G.729} - 2.65603 \times MOS_L^{G.729} \tag{7}$$

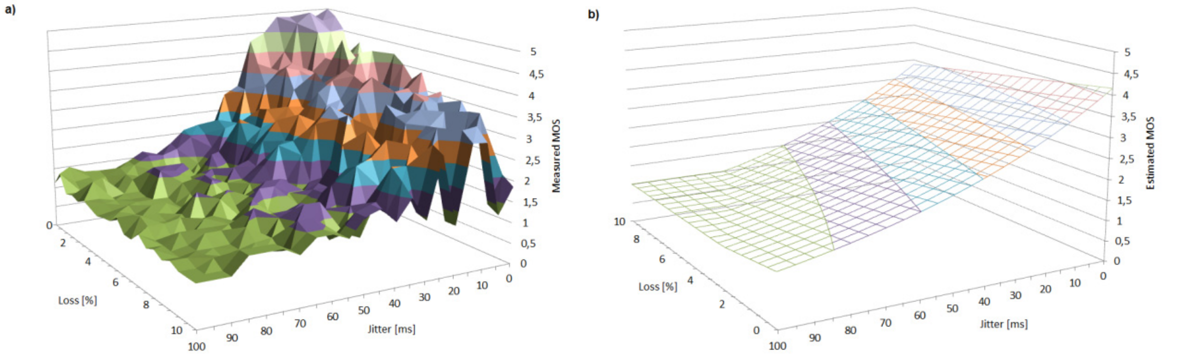

Fig. 4. (a) The impact of jitter and loss on the MOS_M of the G.711; (b) the model for the G.711 MOS_E estimation.

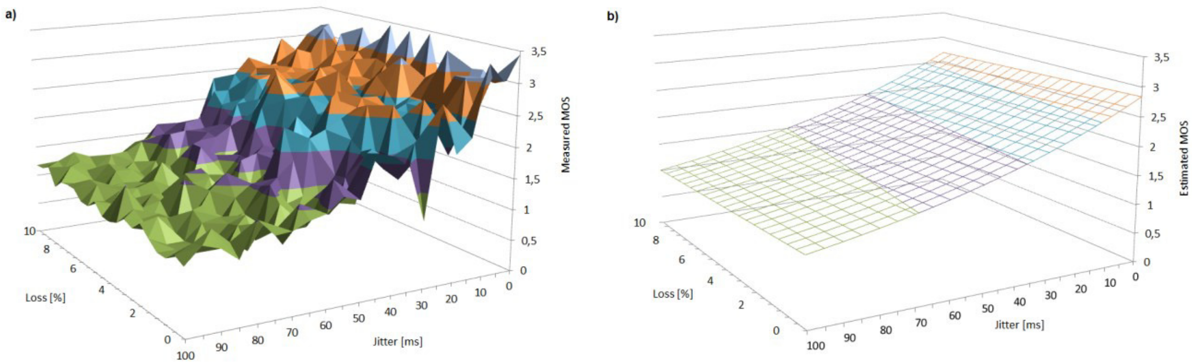

Fig. 5. (a) The impact of jitter and loss on the MOS_M of the G.729; (b) the model for the G.729 MOS_E estimation.

5.3. Test validation

To validate the MOS_E model, the Pearson coefficient (PC), the coefficient of determination (R^2) and standard deviation of regression tests were used. These tests allow us to prove accuracy and usefulness of the model as shown in Tab. 3. The PC of MOS_M and MOS_E values indicates good accuracy by significant linear relationship for both of the VoIP codecs (0.8701 and 0.8029). The determination coefficient shows how well the MOS_M values fit the proposed statistical model (MOS_E values). The model of G.711 has R^2=0.7571, R^2=0.6818 in case of the G.729. The whole model was significantly influenced by relatively high standard deviation of MOS_M.

Table 3. Linear regression details.

Description	MOS_M/MOS_E PC	R^2	σ of regression	σ of MOS_M
G.711	0.8701	0.7571	0.5121	1.0370
G.729	0.8029	0.6818	0.3479	0.6154

To analyze the accuracy of the model, charts representing all the results were traced in detail. As shown in Fig. 6a and Fig. 6b, the ideal distributions between the measured MOS and estimated MOS are represented by red lines. The blue points represent the relationship between the measured (real) MOS and the estimated MOS values obtained by the linear regressions.

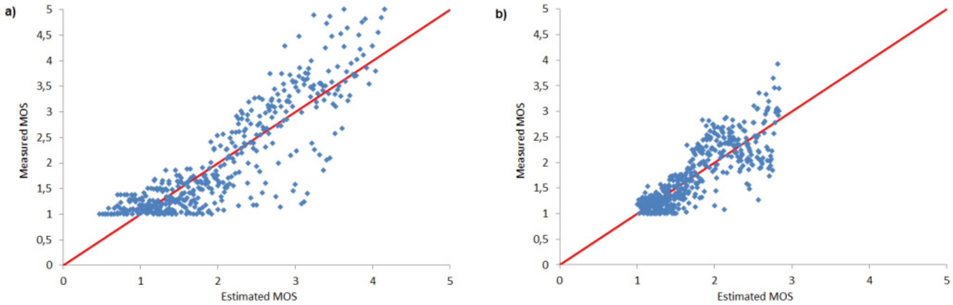

Fig. 6. (a) The accuracy of the G.711 model; (b) the accuracy of the G.729 model.

Nevertheless, these models offer a good way how to estimate the resulting MOS value in real-time. This allows the network administrators and engineers to predict the VoIP quality level that the end-user will give to the content and consequently allows you to avoid the service quality issues.

6. Conclusion

This article deals with the listening quality estimation of VoIP calls in the non-intrusive manner. The impact of delay, jitter and packet loss on VoIP calls was studied. According to measured data a new statistical model suitable for voice listening quality estimation in corporate environments was proposed. In comparison with the existing models, this proposal is based on data measured by AQuA software, a new alternative to PESQ and POLQA algorithms. The model of G.711 quality estimation gives correlation of 87 % between measured and estimated values. The estimated MOS values of G.729 VoIP calls reach correlation of more than 80 % with measured MOS values. Proposed model has a good potential to be considered as the no-reference model used to monitor the listening service quality level as part of the network monitoring systems.

Future work will focus on the remaining QoS parameters influencing the QoE. Besides the examined parameters, packet corruption and packet duplication can be taken into account to improve final accuracy of models. Above that, the impact of other voice codecs quality can be examined.

References

Alvarez, A., Pozueco, L., Cabrero, S., Paneda, X. G., Garcia, R., Melendi, D., Orueta, G. D., 2013. Subjective evaluation of critical success factors for a QoE aware adaptive system. *Computer Communications*. Vol. 36, no. 15–16., pp. 1608-1620. ISSN 0140-3664.

Bayer, N., Xu, B., Rakocevic, V., Habermann, J., 2010. Application-aware scheduling for VoIP in Wireless Mesh Networks. *Computer Networks*. Vol. 54, no. 2, pp. 257–277. ISSN 1389-1286.

Ding, L., Lin, Z., Radwan, A., El-Hennawey, M. S., Goubran, R. A., 2007. Non-intrusive single-ended speech quality assessment in VoIP. *Speech Communication*. Vol. 49, no. 6, pp. 477–489. ISSN 0167-6393.

Hoßfeld, T., Binzenhöfer, A., 2008. Analysis of Skype VoIP traffic in UMTS: End-to-end QoS and QoE measurements. *Computer Networks*. Vol. 52, no. 3, pp. 650–666. ISSN 1389-1286.

Inacio, A., Cruz, R., Nunes, M. S., 2013. Quality user experience in advanced IP video services. *Annals Of Telecommunications*. Vol. 68, no. 3–4, pp. 119–131. ISSN 0003-4347.

ITU-T, 2005. Rec. G.107: The E-model: a computational model for use in transmission planning. [online]. International Telecommunication Union, Geneva. [cited 2014-02-20]. Accessible at: <http://www.itu.int/rec/T-REC-G.107-200503-S/en>.

ITU-T, 1996. Rec. P.800: Methods for subjective determination of transmission quality. [online]. International Telecommunication Union, Geneva. [cited 2014-02-20]. Accessible at: <http://www.itu.int/rec/T-REC-P.800-199608-I/en>.

ITU-T, 2001. Rec. P.862: Perceptual Evaluation of Speech Quality (PESQ): An objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs. [online]. International Telecommunication Union, Geneva. [cited 2014-02-20]. Accessible at: <http://www.itu.int/rec/T-REC-P.862-200102-I/en>.

ITU-T, 2011. Rec. P.863: Perceptual objective listening quality assessment. [online]. International Telecommunication Union, Geneva. [cited 2014-02-20]. Accessible at: <http://www.itu.int/rec/T-REC-P.863-201101-I/en>.

Klein, A., Jakopin. N., 2014. Consumers' willingness-to-pay for mobile telecommunication service bundles. *Telematics and Informatics.* Vol. 31, no. 3, pp. 410-421. ISSN 0736-5853.

Kuipers, F., et al., 2010. Techniques for Measuring Quality of Experience. In: *Proceedings of the 8th International Conference on Wired/Wireless Internet Communications*. pp. 216–227. ISBN 978-3-642-13314-5.

Lozano-Garzon, C., Arriza-Porras, Ch., Rivera-Díaz, S., Riveros-Ardila, H., Donoso, Y., 2012. Mobile Network Qoe/QoS Decision Making tool for Performance Optimization in Critical Web Service. *International Journal of Computers, Communications & Control*. Vol. 7, no. 5, pp. 892–899. ISBN 1841-9836.

Papageorgiou, A., Miede, A., Schullte, S., Schlullter, D., Steinmetz, R., 2013. Decision support for Web service adaptation. *Pervasive and Mobile Computing.* ISSN: 1574-1192.

Santos, C. E. M., Ribeiro, E. P., Pedroso, C. M., 2014. The application of neural networks to improve the quality of experience of video transmission over IP networks. *Engineering Applications of Artificial Intelligence*. Vol. 27, pp. 137–147. ISSN 0952-1976.

Sevana Oy, 2014. *AQuA – Audio Quality Analyzer*. [online]. Sevana Oy. [cited 2014-02-20]. Accessible at: < http://sevana.fi/aqua.php>.

StarTrinity, 2014. StarTrinity SIP Tester. StarTrinity.com [cited 2014-02-20]. Accessible at: <http://startrinity.com/VoIP/SipTester/SipTester.aspx>.

Silva, A. P. C., Varela, M., Silva, E. S., Leao, R. M. M., Rubino, G., 2008. Quality assessment of interactive voice applications. *Computer Networks*. Vol. 52, no. 6, pp. 1179–1192. ISSN 1389-1286.

Stankiewicz, R., Jajszczyk, A., 2011. A survey of QoE assurance in converged networks. *Computer Networks*. Vol. 55, no. 7, pp. 1459–1473. ISSN 1389-1286.

Seppänen, J., Varela, M., Sgora, A., 2014. An autonomous QoE-driven network management framework. *Journal of Visual Communication and Image Representation*. Vol. 25, no. 3, pp. 565–577. ISSN 1047-3203.

Škorpil, V., Šťastný, J., 2006. Back-Propagation and K-Means Algorithms Comparison. In: *8th International Conference on Signal Processing Proceedings*. pp. 1871–1874. ISBN 0-7803-9736-3.

Taboada, I., Liberal, F., Fajardo, J. O., Ayesta, U., 2013. QoE-aware optimalization of multimedia flow scheduling. *Computer Communicatons*. Vol. 36, no. 15–16, pp. 1629–1638. ISSN 0140-3664.